York University at TREC 2011: Medical Records Track

Mariam Daoud, Dawid Kasperowicz, Jun Miao, Jimmy Huang
Information Retrieval and Knowledge Managment Lab
York University, Toronto, Canada
{daoud, dawidk, jhuang}@yorku.ca, jun@cse.yorku.ca

Abstract

In this paper, we present our participation in the Medical Records Track of TREC 2011. The goal of this track is to develop quick search and retrieval tools that are useful for physicians for the purpose to find patients that have similar diseases and/or treatments. To achieve this goal, we propose query expansion and semantic matching models using semantic medical ontologies for medical data retrieval. The query expansion utilizes a medical disease dictionary that presents different possible reformulations given the query disease keywords. For the semantic matching model, we employed BioLabler, a medical annotation tool that allows indexing of queries and documents with UMLS concepts of our choosing. Moreover, the matching model consists of ranking the documents that contain the query concepts according to their scores in the document. We also evaluate a traditional weighting model (BM25), query expansion using relevance feedback under Rocchio's feedback framework and the impact of genre and age filtering, proximity and co-occurrence between disease keywords and procedure/intervention keywords on the retrieval performance.

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**]: H.3.1 Content Analysis and Indexing; H.3.3 Information Search and Retrieval; H.3.4 Systems and Software;

General Terms

Measurement, Performance, Experimentation

Keywords

Disease-based Query Expansion, Semantic Indexing, Conceptual Matching Model, Medical Ontology, Filtering

1 Introduction

Due to the escalating quantity of digitalized medical patient records, the need for advanced information retrieval and knowledge discovery systems increases. Among the applications of computer science in the field of health care, the processing of clinical patient records is playing a increasingly vital role in improving the Electronic Health Record (EHR) systems. An EHR system is defined as a repository of electronically maintained information about a patient's health status and the health care being provided to a patient. In EHR systems, patient data is kept and consists of, but is not limited to: dates and results of screenings, major illnesses and surgeries, in addition to lists of medicines, medicine dosages, allergies, family illnesses, clinical laboratory results, discharge summaries, etc. In the biomedical information retrieval field, the most sought information consists of pattern recognition algorithms, patient clustering procedures, advanced medical information access tools, and recommendations on potential diseases and treatments for patients. Medical pattern recognition consist mainly of detecting changes or regularities within patient's lifestyle and medical data [GAM+05, HHMK10, SJH54]. In [HHMK10], and as such, a time series based method is proposed to analyze and predict personal medical data based on the patients history. In order to explore interrelations among a patient's lifestyle, health statuses were collected for groups of patients. Patient clustering is essential where physicians can assess the effectiveness of previous treatments and recommend new treatments or modifications based on the response of a patient [HAH10]. In [HAH10], medical search and classification tools are proposed for such recommendations. These tools aim at classifying/assigning a patient to predefined disease classes. In biomedical information access, the desired information of a question (query) asked by biologists is usually a list of entities covering different aspects that are related to the question [HCRR07]. In a general context, the various aspects could cover disease, symptoms, treatment, etc, thus; it is important for a biomedical IR system to be able to provide comprehensive and diverse answers to fulfill biologists' information needs. In [YHL10], a re-ranking method to promote ranking diversity for biomedical information retrieval is proposed where aspects covered by retrieved passages are detected and explicitly presented by Wikipedia concepts.

The goal of the Medical Records Track in TREC 2011 is to foster research on providing content-based access to the free-text fields of electronic medical records. A pre-defined set of topics specify a particular disease/condition set and a particular treatment/intervention set. For example, a topic might be as follows: "find patients with gastroesophageal reflux disease who had an upper endoscopy". Thus, for a given query, the system should return a list of visits, ranked in decreasing likelihood that a patient's visit satisfies the given queries specifications.

In the context of medical search, traditional information retrieval has the following major limitations [ZH06]: the frequent use of (possibly non-standardized) acronyms the presence of homonyms (the same word referring to two or more different entities) and synonyms (two or more words referring to the same entity). These limitations are mainly due to the keyword-based matching model in traditional IR since a document is considered a bag of words, while a medical entity is usually represented by a set of dependent and adjacent terms. In order to alleviate such limitations, a challenging task in clinical text retrieval is to find accurate synonyms or name variants for medical entities and unify their representation for both queries and documents.

With respect to these challenges, our contributions in this paper concern semantic query expansion and conceptual matching models using semantic resources. For query expansion, we use medical disease vocabulary that provides various possible synonym reformulations of a disease keyword given in the query. For the conceptual matching model, we use BioLabler, a medical annotation tool that allows indexing queries and documents with Unified Medical Language System (UMLS) concepts of our choosing, and rank the documents using their conceptual representation. We also evaluate a query expansion technique using Rocchio algorithm and the impact of age and gender filters, along with the proximity and co-occurrence between disease keywords, and procedure/intervention keywords on the retrieval performance.

This paper is organized as follows: Section 2 presents a query expansion technique using Rocchio's algorithm with gender-age filtering. Section 3 details our semantic query expansion technique using disease synonym dictionary. Section 4 illustrates our semantic matching model based on conceptual query and document indexing using UMLS. Section 5 presents our experiments and runs. In section 6, we conclude the paper and present our future work.

2 Query Expansion using Rocchio with gender-age Filtering

One of our contributions is to perform a query expansion technique using Rocchio's algorithm coupled with gender and age filtering. The Rocchio relevance feedback [Roc71] is a classic algorithm for implementing relevance feedback. It models a way of incorporating relevance feedback information into the vector space model. In particular, it take a set of documents for feedback and the candidate terms in this set of documents are ranked according to the following formula:

$$Q_1 = \alpha * Q_0 + \beta * \sum_{rel} \frac{D_i}{|D_i|} \tag{1}$$

where Q_0 and Q_1 represent the initial and first iteration query vectors, D_i represents document weight vectors, $|D_i|$ is the corresponding Euclidian vector length, and α, β are tuning constants.

After performing query expansion, we generate a list of patient reports that is then converted to a visit list where the score of each visit is computed as the sum of the patient report scores. Based on the visit result list, we filter the results according to gender and age constraints identified in the query keywords. Topics like: "Women with osteopenia" or "Adults who received a coronary stent during an admission", contain either gender or age constraints for finding the relevant patient visit. For this purpose, we extract age and gender information from each report in the dataset, based on finding specific gender-related or age-related keywords such as "age[in xxs]", "lady", "male", etc. Gender and age information is then associated with the patient's visit and all the patient's reports that are associated with this visit, since the age and gender elements are unique for a visit. Finally, we remove all visits that do not satisfy the age and gender query requirements to obtain our final results.

3 Semantic Query Expansion Using Disease Synonyms

Our query expansion technique focuses on adding medical disease keywords to the users query where keywords are added from a medical synonym dictionary, namely Polysearch[1]. The main motivation for using this technique is that clinical reports commonly use abbreviations and synonyms, which must be accounted for in order to get optimal results. Algorithm 1 presents our query expansion technique using a medical dictionary.

We first manually identify disease keywords from the topic and then automatically extract associated synonym from the medical disease dictionary based on maximum keyword overlap. Identifying the disease keywords in the query could be made automatically by using a medical annotation tool, an area left for future work. Each entry in the dictionary contains a synonym set identifier and a synonym set that contains different possible reformulations of the same disease. Based on keyword matching, we obtain a set of synonyms for each topic and each set represents a single disease. We only keep the sets that have the maximum amount of overlap with the topic at hand. The new topic is then formed by appending the original topic with the terms of all the candidate synonym sets. Topic 102 is an example where an abbreviation of a disease is used, namely GERD :"Patients with complicated GERD who receive endoscopy". The corresponding description for "GERD" is "gastroesophageal reflux disease". The synonym set identified for topic 102 is as follows: "DID64117, gastroesophageal reflux disease, Esophageal reflux, Oesophageal reflux, Gastroesophageal reflux disease, GERD, GE REFLUX".

[1] $http://wishart.biology.ualberta.ca/polysearch/include/disease_IDlist.txt$

Algorithm 1 Selecting disease synonym set for semantic query expansion

INPUT: query disease keywords Q_0,
OUTPUT: Synonym set list,
// *Extract the synonym set candidates from the dictionary*
Let S be the list of the synonym sets in the disease dictionary
$S = S_1, S_2, S_3, ..., S_i, ..S_n$
for each synonym set S_i in S **do**
  // *compute the overlap with the query disease terms*
  for each synonym set s_{ij} in S_i **do**
    $Overlap(s_{ij}, q) = CommonTerms(s_{ij}, q)$
  end for
  $Overlap(S_i, q) = Argmax_j Overlap(s_{ij}, q)$
  //Select the synonym sets that have the maximum overlap with the query
  if($Overlap(S_i, q) = |q|)Then SynList = SynList \bigcup S_i$
end for
// Append the query with the synonym set terms
$Q_n = Q_0 + \bigcup_{j/S_i \in SynList \wedge s_{ij} \in S_i} t \in s_{ij}$

4 Semantic Matching Model Using UMLS Conceptual index

Our second contribution consists of conceptual query and document indexing using UMLS, and a conceptual matching model based on concept confidence weights in the document. Using semantic indexing, different free text reformulations or synonyms of a medical entity are mapped to the same medical concepts in a predefined ontology. Acronyms and anonyms are likewise mapped to medical concepts in the ontology so that their representation in the dataset is unified.

The conceptual indexing of queries and documents is based on using UMLS concepts. We use an online biomedical text mining tool named BioLabeler[2] which associates UMLS concepts for any given text. In order to satisfy the query requirements for finding patients that have both specific diseases and procedures, two indexes are created: a disease-based index and a procedure-based index. All UMLS sources obtained by BioLabler are used to generate the two indexes for queries and patient reports. Each medical concept contains statistical information to determine the relevance of a particular medical concept to the patient report, specifically its normalized weight and regular weight. This weight reflects the degree of relevancy for a particular concept. For Topic 101 : “Patients with hearing loss”, a high ranked disease concept according to BioLabler is: “Hearing Loss, Central#Central hearing loss#Central hearing loss#Central hearing loss#Central Hearing Loss” with a concept unique identifier (CUI) of C0018776. A candidate procedural concept reported by BioLabler is “Audiometry#Audiometry#Audiometry#Audiometry#Audiometry” with a CUI of C0004286. Once queries and documents are indexed at a conceptual level, we use a term-based matching model that is based on matching terms of query concepts with terms of the document concepts in order to compute a relevance score of each document with respect to the query. Given the disease-based query concept vector $Q_d = (C_1, C_2, .., C_m)$ and the disease-based document concept vector $D_d = (C_1, C_2, .., C_n)$, the conceptual score of the document is computed by matching the documents that contain concepts whose terms overlap with the query concept terms and ranking them according to the overlapping concept weights in the document as follows:

$$Score_d(D) = \sum_{C_i \in Q_d} \frac{1}{n} \sum_{C_j \in D_d / \exists t \in Terms(C_i) \cap Terms(C_j)} w(C_j, D_d) \qquad (2)$$

where $Score_d(D)$ is the conceptual score of the document obtained using disease-based index, C_i is a concept in query Q_d and C_j is a concept in document D_d, $Terms(C_i)$ and $Terms(C_j)$ are the

[2] http://www.biolabeler.com/bioLabeler/

lists of terms associated with concepts C_i and C_j respectively. $w(C_i, Q_d)$ and $w(C_j, D_d)$ represent the weight of concept C_i in query Q_d and the weight of concept C_j in document D_d respectively. Based on a procedure-based index, we calculate $Score_p(D)$ using the same formula.

Another way for computing a pure conceptual score of the document is to use the dot product formula between concepts of the query and concepts of the document as follows:

$$Score_d(D) = \sum_{C_i \in Q_d} w(C_i, Q_d) * w(C_i, D_d) \tag{3}$$

where $Score_d(D)$ is the conceptual score of the document obtained using disease-based index, $w(C_i, Q_d)$ and $w(C_i, D_d)$ represent the weight of concept C_i in query Q_d and document D_d respectively. Based on a procedure-based index, we calculate $Score_p(D)$ using the same formula.

The final score of the document is based on combining the disease-based document score and the procedure-based score as follows:

$$Score_f(D) = Score_d(D) + Score_p(D) \tag{4}$$

5 Experiments

All the topics consist of finding patients who have specific diseases, and/or have received specific treatments or interventions. The testing document collection for the Medical Records Track is a set of de-identified medical records made available for research use through the University of Pittsburgh BLULab NLP Repository. Each medical record is associated with a unique patient's visit. The Medical Records track use the visit as the response unit. Thus, the retrieval system must return visitIDs, where relevance judgments are based on the visit as a whole.

5.1 Runs

Our main goals for the runs are the following: (1) Observing the impact of Rocchio's query expansion with and without applying the filters, (2) Evaluating the impact of disease synonym query expansion, (3) Evaluating the impact of semantic indexing and conceptual matching and, (4) Evaluating the impact of proximity and co-occurrence between the disease and treatment/intervention on the retrieval performance. Table 1 presents the four official submitted runs and eight additional runs.

Table 1: Submitted runs

Official Runs	**Description**
Baseline	Baseline, plain terms with BM25.
QE-Filter	Query expansion using Rocchio + age and gender filtering.
QE-DisSyn	Query expansion using disease synonym dictionary.
CM-UMLS	Term-based conceptual matching using UMLS concepts.
Other runs	**Description**
QE	Query expansion using Rocchio algorithm.
Prox	Proximity run between disease keywords.
Cooc	Co-occurrence run between disease and treatment keywords.
Prox-Cooc	Combining proximity between disease keywords and their co-occurrence with treatment keywords.
CMC-UMLS	Concept-based matching using UMLS concepts.
CMC-MSH1	Concept-based matching using only Mesh ontology and the most relevant concept of the query.
CMC-MSH5	Concept-based matching using only Mesh ontology and the top 5 relevant concepts of the query.

We use the Terrier search engine[3] where Porter stemming and stopword removal is conducted in indexing and searching processes for our runs. The baseline run consists of a probabilistic retrieval model based on BM25 scoring function where we set b to 0.75, k1 to 1.2 and k3 to 8. For the QE run, we use the Dirichlet language model with a miu of 2500, 3 feedback documents and 10 feedback terms. For the Prox run, we consider the proximity between disease keywords in the queries by using quotation marks on disease keywords as a query operator that is built into the Terrier system. For the Cooc run, we consider the co-occurrence between the disease keywords, and the treatment/intervention keywords by using "+" symbol as a query operator that is built into Terrier. The Prox- cooc run considers both the proximity between disease keywords and their co-occurrence with treatment/intervention if it exists in the query. For the Prox, Cooc and Prox-cooc runs, we manually add the query operators, namely the quotation marks and "+" symbol, and if no results are retrieved for a particular topic, the baseline run is associated for said topic. CM-UMLS run is performed using Formula 2. CMC-UMLS, CMC-MSH1 and CMC-MSH5 runs are performed using Formula 3.

5.2 Results

Table 2 summarizes the results of our official submitted runs and the additional runs that were conducted according to the following official measures: bpref, R-prec and P10. Figure 3 details the P10 values per topic of each run. We observe that the best among the official submitted runs is the baseline run.

The overall values of the official measures between QE-Filter and the baseline runs do not look significantly different, which leads us to conclude that query expansion used with gender and age filtering do not have a significant impact on the retrieval performance. Using only query expansion achieves better performance than the baseline in terms of bpref and R-prec but not at P10. This illustrates that query expansion alone has a slightly positive impact on the performance while the filters have a negative impact on the performance. The negative impact of the filtering process might be due to the certainty/uncertainty that lead to assign the same gender and age information to all the reports of the same visit, based on one document's annotation. Conversely, out of 100865 reports in the collection, 89669 reports are associated to adult patients of more than 20 years of age, and 8979 reports do not have age signatures. This reduce the impact of removing reports from the small set of child patients (patients under the age of 20).

Table 2: Results

Official runs	**bpref**	**R-prec**	**P10**
Baseline	0.3632	0.2606	0.4176
QE-Filter	0.3500	0.2605	0.4147
QE-DisSyn	0.3399	0.1927	0.3353
CM-UMLS	0.0834	0.0159	0.0382
Other runs	**bpref**	**R-prec**	**P10**
QE	0.3778 (4.02%)	0.2663 (2.19%)	0.4086 (-2.16%)
Prox	0.3668 (1.00%)	0.2897 (11.17%)	0.4735 (13.39%)
Cooc	0.4164 (4.65%)	0.3262 (25.17%)	0.5059 (21.14%)
Prox-cooc	0.3730 (2.70%)	0.3006 (15.34%)	0.5088 (21.84%)
CMC-UMLS	0.1742	0.06	0.1147
CMC-MSH1	0.2528	0.1725	0.2559
CMC-MSH5	0.2817	0.1639	0.2143

The QE-DisSyn run also has a negative impact on the system's performance, possibly due to the use of synonym terms as a bag of words to form a new query. A disease name consists usually of a sequence of terms that has a meaning with a specific term order. For instance "Hepatitis C"

[3] www.terrier.org

is an example of such a case. In addition, a single term disease name is usually ambiguous and could be associated to irrelevant synonym sets of the dictionary.

The CM-UMLS run achieves the lowest negative performance. The main reason is related to the concept extraction accuracy provided by BioLabler annotation tool, and the limitations of term-based matching between concepts of the query and concepts of the document.

By observing the Prox, Cooc and Prox-cooc runs, we can clearly conclude that there is a positive impact of using proximity between the disease keywords in the query and the co-occurrence between disease and treatment/intervention keywords. Our first observation is that there is a small improvement at bpref than R-prec and P10, primarily due to the robustness of bpref measure in the face of incomplete relevance information where the scores for R-precision, and P10 are completely determined by the ranks of the relevant documents in the result set. These measures make no distinction in pooled collections between documents that are explicitly judged as irrelevant and documents that are assumed to be irrelevant because they are not judged [BV04].

The Prox run achieves a positive improvement at R-prec and P10. This proves to be a more meaningful representation of the topic, where disease keywords reflect the meaning of the disease concepts instead of a bag of words. The Co-occurrence run is the best run at all precision measures. This is mainly due to satisfying the co-occurrence of both the disease and the treatment/intervention in a patient report with respect to the topic specifications. This run achieves better performance than Prox run since it forces the retrieval of only the reports that contain both disease and treatment while the proximity run could retrieve reports containing only a disease or treatment/intervention. Combining both proximity with co-occurrence query operators boost the performance at P10 over the Prox and Cooc runs. The negative impact of using both proximity and co-occurrence is the limited set of retrieved results for the topics, since the retrieval process is based on removing out the reports that do not satisfy the query operators (either proximity or co-occurrence) and the exact keyword matching schema between queries and reports. A better exploitation of both features is to re-rank a baseline run based on classic term-based matching in order to get a rich result set.

For the concept-based matching runs, CMC-UMLS has superior performance over CM-UMLS due to the concept-based matching being more efficient than the term-based matching that is applied over the concept terms. More interestingly, CMC-MSH1 and CMC-MSH5 performance is superior to CM-UMLS and CMC-UMLS, which proves that using only the MESH source for indexing and retrieval achieves the highest precision values for all measures, compared to other conceptual matching runs. However, no run outperforms the baseline run due to the low precision of concept-based indexing for both queries and documents. In the future, we plan to apply a disambiguation technique for improving the indexing step using medical concepts.

In order to better understand the performance of the different models, we present in Table 3 the P10 values for each topic of the baseline, QE-filter, QE-syn and Cooc runs.

We observe a performance variability across topics where some topics have a low precision at P10 in the baseline. For instance, topic 124 and topic 125 have a null value for P10 for all the runs. Topic 124 addresses patients with episodes of acute loss of vision secondary to glaucoma where there are only 6 relevant visits for this topic. Topic 125 addresses patients who are infected with both "hepatitis C" and HIV, where there are only 2 documents that contain both terms is 2 and the number of relevant documents is 14. The main challenge in this topic is the use of the acronym HIV and the bag of words representation of Hepatitis C. Query expansion using synonym has no effect on this topic since the synonym terms are added as a bag of words, where adding terms as a bag of words only allows handling of acronym substitution.

QE-filter run has improved only 6 queries where the filters have an impact on only 5 queries (topic 109, 112, 113, 114, 115, 118, 119), which address gender and age constraints. Among 5 topics, only 2 topics have been improved. The fact that the filters have no effect on the rest of topics could be due to assigning the same gender and age information to all the reports of the same visit, based on one document's annotation.

The QE-syn run has improved only 4 topics. This could be related to adding noisy terms to the original topic or due to the limited impact of appending the query with a bag of words term list. Cooc run has improved 15 topics where the use of co-occurrence between the disease and treatment/intervention keywords has no effect on the retrieval performance for the rest of topics. This could be due to the use of different lexical names of the same disease or treatment in the patient report where the Cooc run cannot handle this issue since it is based on retrieving documents that contain the exact terms of the topic.

Table 3: P10 values for each topic of the official runs

Topic	**baseline**	**QE-filter**	**QE-syn**	**Cooc**	**Topic**	**baseline**	**QE-filter**	**QE-syn**	**Cooc**
101	0.20	**0.50+**	**0.30+**	**0.50+**	118	0.10	0.10	0	0.10
102	0.40	0.40	0.30	**0.50+**	119	0.90	0.90	0.40	0.90
103	0.20	0.20	0.10	0.20	120	0.70	0.60	0.70	**0.80+**
104	0.10	0.10	0.10	0.10	121	0.30	0.30	0	0.20
105	1	1	0.90	1.	122	0.10	0.10	0.10	**0.60+**
106	0.30	0.10	**0.90+**	**0.70+**	123	0.30	0.10	0	**0.90+**
107	0.30	**0.40+**	**0.40+**	**0.90+**	124	0	0	0	0
108	0.10	0	0	**0.20+**	125	0	0	0	0
109	0.90	**1+**	0.90	0.90	126	0.20	0.20	0	0.20
110	0.70	0.60	0.60	**0.80+**	127	0.80	0.80	0.50	0.70
111	0.20	0.10	0	0.20	128	0.80	0.70	0.70	0
112	0.90	0.80	0.80	0.90	129	0.20	0.10	0	0.10
113	0.60	0.60	0.40	**0.80+**	131	0.60	**0.70+**	0.40	0.40
114	0.60	**0.80+**	0.60	**0.80+**	132	0.90	0.90	0.90	0.90
115	0.40	0.30	0.30	**0.60+**	133	0.10	0.10	0.10	0
116	0.30	**0.40+**	0	**0.40+**	134	0.30	0.30	0.20	0.10
117	0	0	0	**0.90+**	135	0.70	**0.90+**	**0.80+**	**0.90+**

6 Conclusions

The TREC Medical Records Track presents a challenging ad-hoc retrieval task where the focus is on the reasoning part of the system. In this paper, we present our participation in the Medical Records Track of TREC 2011. First, we evaluate traditional weighting models (BM25) and query expansion using Rocchio's framework. Second, we evaluate our proposed methods for medical retrieval, mainly age and gender filtering, semantic query expansion using disease synonym, and conceptual weighting model. We also evaluate the impact of proximity and co-occurrence on the disease and treatment/intervention query keywords. We conclude that the lack of significant improvement on the overall retrieval results using synonym-based reformulated queries show that method has limited semantic capabilities where a semantic representation of disease and treatment needs further improvements. At conceptual level, using BioLabler for indexing has also shown limited impact on the retrieval performance, where a disambiguation technique should be used to exclude irrelevant concepts from both query and document representation. A more elaborated conceptual matching model is also needed where appropriate concepts weights could contribute better for improving the performance. Our final conclusion concerns the positive impact of using proximity and co-occurrence features to satisfy the user's information need.

Future work will focus on exploring different strategies for identifying the most relevant disease synonym to append the query where we plan to consider term correlations in representing disease and treatment/intervention concepts. Also, we plan to apply disambiguation techniques to increase the accuracy of the conceptual query and document representation in order to achieve better conceptual matching model performance.

7 Acknowledgements

This research is jointly supported by NSERC of Canada and the Early Researcher/Premier's Research Excellence Award. We would like to thank Robert Saggiorato, Atanaz Babashzadeh, in addition to Jose Luis Marina and Csar Silgo from the BioLabeler team, for their valuable help while conducting our work.

References

[BV04] Chris Buckley and Ellen M. Voorhees. Retrieval evaluation with incomplete information. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '04, pages 25–32, New York, NY, USA, 2004. ACM.

[GAM+05] Amit X. Garg, Neill K. J. Adhikari, Heather McDonald, M. Patricia Rosas-Arellano, P. J. Devereaux, Joseph Beyene, Justina Sam, and R. Brian Haynes. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: A systematic review. *Jama*, 293(10), 2005.

[HAH10] Xiangji Huang, Aijun An, and Qinmin Hu. Medical search and classification tools for recommendation. In Fabio Crestani, Stéphane Marchand-Maillet, Hsin-Hsi Chen, Efthimis N. Efthimiadis, and Jacques Savoy, editors, *SIGIR*, page 707. ACM, 2010.

[HCRR07] William R. Hersh, Aaron M. Cohen, Lynn Ruslen, and Phoebe M. Roberts. Trec 2007 genomics track overview. In Ellen M. Voorhees and Lori P. Buckland, editors, *TREC*, volume Special Publication 500-274. National Institute of Standards and Technology (NIST), 2007.

[HHMK10] Qinwin Vivian Hu, Jimmy Xiangji Huang, William Melek, and C. Joseph Kurian. A time series based method for analyzing and predicting personalized medical data. In Yiyu Yao, Ron Sun, Tomaso Poggio, Jiming Liu, Ning Zhong, and Jimmy Huang, editors, *Brain Informatics*, volume 6334 of *Lecture Notes in Computer Science*, pages 288–298. Springer, 2010.

[Roc71] J. Rocchio. *Relevance Feedback in Information Retrieval*, pages 313–323. 1971.

[SJH54] William W Stead, Leland E Garrett Jr, and William E Hammond. Practicing nephrology with a computerized medical record. *Kidney Int.*, 24(4), 446–454.

[YHL10] Xiaoshi Yin, Xiangji Huang, and Zhoujun Li. Promoting ranking diversity for biomedical information retrieval using wikipedia. In Cathal Gurrin, Yulan He, Gabriella Kazai, Udo Kruschwitz, Suzanne Little, Thomas Roelleke, Stefan M. Rüger, and Keith van Rijsbergen, editors, *ECIR*, volume 5993 of *Lecture Notes in Computer Science*, pages 495–507. Springer, 2010.